AD-A056 949 WASHINGTON UNIV SEATTLE DEPT OF PSYCHOLOGY F/G 5/1

THE EQUIVALENCE INTERVAL AS A MEASURE OF UNCERTAINTY. (U)

MAR 78 J R LARSON, A M REENAN N00014-76-C-0193

UNCLASSIFIED TR-77-9 NL

1 OF 1

AD A056949

END

DATE FILMED

9 -78

DDC

LEVEL II

11

SC

AD A056949

DECISION MAKING RESEARCH

DEPARTMENT OF PSYCHOLOGY
DEPARTMENT OF MANAGEMENT AND ORGANIZATION
UNIVERSITY OF WASHINGTON, SEATTLE, WASHINGTON

DECISION MAKING RESEARCH
DEPARTMENT OF PSYCHOLOGY
UNIVERSITY OF WASHINGTON
SEATTLE, WASHINGTON

LEVEL II

11

AD A056949

AD No.
DDC FILE COPY

THE EQUIVALANCE INTERVAL AS A[1] MEASURE OF UNCERTAINTY

James R. Larson, Jr. and
Andrew M. Reenan

University of Washington
Seattle, Washington

Technical Report 77-9

March 1978

DISTRIBUTION IS UNLIMITED

Office of Naval Research Contract N00014-76-C-0193

(Terence R. Mitchell and Lee Roy Beach, Investigators)

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER: (14) TR-77-9 / 2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle): (6) The Equivalance Interval as a Measure of Uncertainty.	5. TYPE OF REPORT & PERIOD COVERED: (9) Technical Report
	6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s): (10) James R. Larson, Jr.; Andrew M. Reenan	8. CONTRACT OR GRANT NUMBER(s): (15) N00014-76-C-0193
9. PERFORMING ORGANIZATION NAME AND ADDRESS: Decision Making Research, Department of Psychology NI-25, University of Washington, Seattle, WA 98195	10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS: (11)
11. CONTROLLING OFFICE NAME AND ADDRESS: Organizational Effectiveness Research Programs, Office of Naval Research (Code 452), Arlington, VA 22217	12. REPORT DATE: March 1978
	13. NUMBER OF PAGES: 18
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office): (12) 25p.	15. SECURITY CLASS. (of this report): UNCLASSIFIED
	15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Equivalence Interval (EI)	Non-Confidence Measure
Uncertainty Measure	Judgment Uncertainty
Judgment Task	Judgment Accuracy
Confidence Interval	Standard of Comparison

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

Subjects made a series of 15 judgments, one about each of five stimuli in three different stimulus classes. Then, supposing that their answers were not exactly correct, half of the subjects were asked to indicate the range of values within which they were reasonably certain that the correct answer did in fact lie. This range is termed an equivalence interval. The remaining subjects were simply asked to rate their confidence in the accuracy of each judgment. It was found that the width of the equivalence intervals correlated

(over)

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE
S/N 0102-LF-014-6601

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

384783

20. quite highly with the confidence ratings: As subjects in one group became less confident in the accuracy of their judgments the equivalence intervals given by subjects in the other group became wider. In addition, both measures indicated that the subjects were most uncertain about their judgments when clearly defined standards of comparison were least available. The results were discussed in terms of the usefulness of the equivalence interval technique as a measure of uncertainty.

ACCESSION for

NTIS White Section ☒
DDC Buff Section ☐
UNANNOUNCED ☐
JUSTIFICATION..........

BY..........
DISTRIBUTION/AVAILABILITY CODES

Dist.	AVAIL. and/or SPECIAL	
A		

The Equivalence Interval as a Measure of Uncertainty[1]

The concept of uncertainty has become increasingly important for understanding the decisions people make in a wide variety of situations: from organizational level decisions that have significant impact upon an organization's effectiveness (e.g. Downey & Slocum, 1975; Lawrence & Lorsch, 1967; Thompson, 1967) to fairly mundane sorts of individual level decisions, such as selecting among bets (e.g. Ellsberg, 1961; Raiffa, 1961) and employment conditions (e.g. Larson, Note 1; Larson & Mitchell, 1977) in contrived experimental settings. Yet, relatively little empirical work has been done to delineate either the personal or situational determinants of uncertainty. Moreover, the research that has been done has relied on criterion measures of uncertainty that no only are applicable in just a few settings, but that also tend to have rather low reliability and validity (e.g. Downey & Slocum, 1975; Downey, Hellriegel & Slocum, 1975). The primary purpose of the present study was to investigate the usefulness of a new measure of uncertainty which may be applied in a wide variety of settings.

Uncertainty can be defined as a subjective state in which individuals feel unable to make precise judgments about some characteristic of a given entity, situation, relationship, or event (Larson & Mitchell, 1977). The less precise the judgments the more uncertain the individuals are about the characteristic in question. Defined in this manner, uncertainty is closely related to confidence in the accuracy of a judgment: As uncertainty increases, confidence in the accuracy of a judgment should decrease.

Using this definition, it seems reasonable to measure uncertainty by structuring the judgment task to allow individuals to respond in either more or less precise terms. A measurement technique used to investigate ranges of subjectively acceptable error is well adapted to this purpose (e.g. Beach et al., 1974; Beach & Solak, 1969; Laestadius, 1970). Respondents are first asked to make judgments about some quantitation characteristic of a given stimulus (e.g. weight, size, net earnings, etc.). Then, supposing that their answers are not exactly correct, they are asked to go back and indicate the range of possible values of the characteristic in question that the stimulus could have and still leave them confident that their original judgment was essentially correct, or "in the ballpark." This range of values is termed an <u>equivalence interval</u>, since it is assumed that all of the values falling within it are perceived by the respondents as essentially equivalent to their initial judgment in terms of accuracy. With regard to the present discussion, the size of the equivalence interval can be taken as a measure of uncertainty: The more uncertain individuals are about the correctness of their judgments the larger should be the size of their equivalence intervals. Therefore, it was hypothesized that the width of the equivalence interval will be highly correlated with a rating of confidence in the accuracy of a judgment.

A secondary purpose of the present study was to investigate the effect of having readily available standards of comparison on judgment uncertainty. The judgment process is by its very nature a comparative one in which the characteristic to be judged is compared to some known standard or anchor point (c.f. Stevens, 1966). It was hypothesized that as the actual value of the characteristic to be judged approaches a clearly defined anchor point, judgments about that value will become easier to make, and individuals will therefore tend to be less uncertain about their accuracy. Conversely, individuals will in general be more uncertain about the accuracy of such judgments the further the actual value of the characteristic in question is from a clearly defined anchor point. This hypothesis was tested by assessing subjects' uncertainty about judgments involving stimuli that were either near to or far from a clearly defined anchor point.

Method

Overview

Subjects were asked to make a series of 15 judgments, one about each of five stimuli in three different stimulus classes. Within each class the stimuli varied in the extent to which they were near to or far from either the maximum or minimum possible value of that stimulus. Steps were taken to establish the maximum and minimum possible values of each stimulus class as clearly defined anchor points. After making each judgment the subjects were asked to use one of two methods to indicate how uncertain they were about the accuracy of that judgment. Half of the subjects used a separate bi-polar rating scale to indicate their uncertainty, while the other half constructed equivalence intervals.

Subjects

Sixty undergraduate students enrolled in lower division psychology courses at the University of Washington participated in the study one at a time. They each received one half hour of experimental credit for participating.

Judgment Task

The subjects were required to make five separate judgments of fullness, numerocity, and time. A different type of stimulus was used for each type of judgment.

Fullness. The first set of stimuli consisted of five small sealed opaque paper cartons of uniform size and shape. Each carton held a different number of marbles. The subjects were required to examine each carton and estimate the number of marbles it held. The cartons held 6, 24, 42, 60, and 78 marbles respectively. The subjects were informed that a maximum of 85 marbles could be fit into any one carton. They were asked to indicate their best guess about the exact number of marbles in each carton by placing an "X" at the appropriate point on a numberline marked with 86 points, from 0 to 85.

<u>Numerocity</u>. The second set of stimuli consisted of five slides of 100 red and blue disks intermixed randomly in a 10 x 10 matrix. The slides were presented tachistoscopically one at a time for approximately .25 seconds. Each slide showed a different number of red and blue disks. The subjects were required to estimate the number of red disks pictured in each. The five slides contained 8, 29, 50, 71, and 92 red disks respectively. The maximum possible number of red disks was 100. The subjects were asked to indicate their best guess about the exact number of red disks pictured in each slide by placing an "X" at the appropriate point on a numberline ranging from 0 to 100.

<u>Time</u>. The final set of stimuli consisted of five time intervals: 6, 18, 30, 42, and 54 seconds. The subjects were asked to estimate the length of time that elapsed between two signals given by the experimenter. To prevent them from counting or using some other method to record the passage of time, the subjects were required to read a long series of three-letter nonsense syllables presented individually on index cards during each interval. The subjects were told that no interval would be longer than 60 seconds. Again, they were asked to indicate their best guess about the exact length of each time interval by placing an "X" at the appropriate point on a numberline ranging from 0 to 60 seconds.

Procedure

The experimenter began by describing the purpose of the study, stating that people's accuracy in judging various characteristics of a wide variety of stimuli was being investigated. The fullness, numerocity, and time judgment tasks were then described. When the subjects indicated that they understood what they were to do, they were presented with the first set of stimuli. For each set of stimuli the subjects were first given two standards representing the maximum and minimum possible values of the stimulus class. Thus, for example, before making the fullness judgment the subjects were given two cartons identical to those about which they had to make a judgment. One of these cartons was completely empty, while the other held 85 marbles, the maximum number that it could possibly hold. These two cartons were clearly labeled with the number of marbles they held. The subjects were encouraged to use these as standards of comparison when making their estimates about the number of marbles in each of the unknown cartons. Similarly, for the numerocity judgments two labeled slides, one composed of 100 red disks and the other composed of 100 blue disks, were presented before the set unknown slides were presented. For the time judgment the subjects were given two initial practice trials lasting for 60 seconds each. The length of these two practice trials was clearly stated by the experimenter both before and after they occurred. Each subject made the fullness judgments first, followed by the numerocity judgments, and then the time judgments. However, the five stimuli within each judgment type were presented in different random orders.

Dependent Measures

Two different measures of uncertainty about the accuracy of each judgment were obtained. All of the subjects first reported their best guess about the exact value of each stimulus. Then, half of the subjects were asked to construct an equivalence interval around this guess in the manner described above: They indicated the range of values, both above and below their guess, within which they were reasonably certain that the correct answer would lie and outside of which they were reasonably certain that the correct answer did not lie.

The remaining half of the subjects were asked to report how confident they were about the accuracy of each judgment by placing an "X" at the appropriate point on a separate 21-point bi-polar adjective scale ranging from "quite confident" to "not at all confident." For ease of comparison with the equivalence interval measure, the responses to the confidence ratings were scored so that "quite confident" received a value of 0 and "not at all confident" received a value of 20. Scores computed in this way thus reflect the subjects' degree of "non-confidence."

Results

Judgment Uncertainty

Figure 1 shows the mean uncertainty ratings for each judgment using both the equivalence interval measure and the non-confidence measure. Repeated measures analyses of variance were computed for each judgment type using the equivalence interval measure and the non-confidence measure separately as dependent variables. The F-ratios from these analyses are presented in Table 1.

Insert Figure 1 and Table 1 about here

As can be seen in Figure 1, there was a high degree of similarity in the overall pattern of means for the two different measures. The correlation

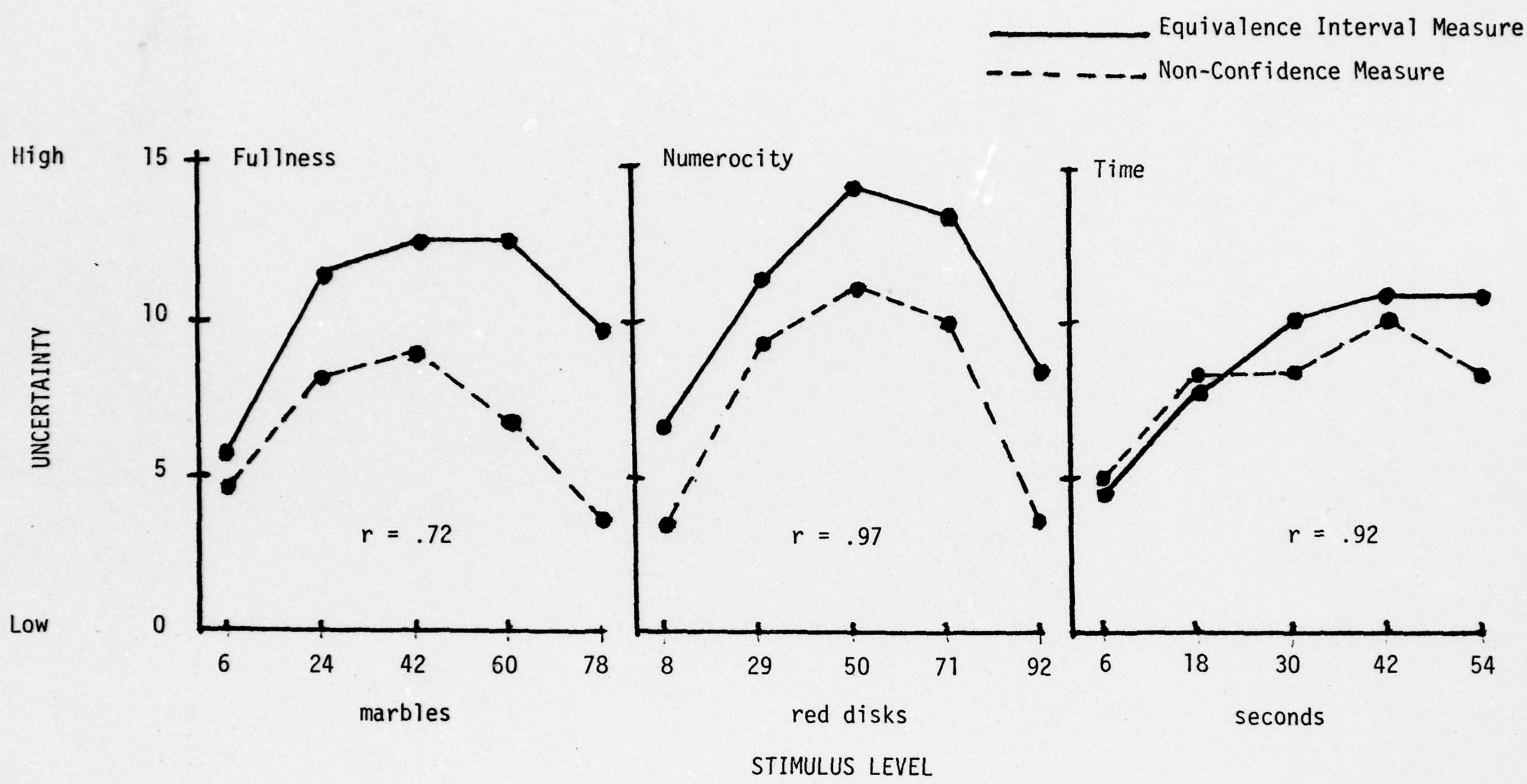

Figure 1. Mean uncertainty ratings using the equivalence interval measure and the non-confidence measure.

Table 1

F-Ratios from the Analyses of Variance Using the Equivalence Interval Measure and the Non-Confidence Measure

Judgment Type	F-Ratios Overall[a]	Linear[b]	Quadratic[b]
	Equivalence Interval Measure		
Fullness	20.25***	21.19***	57.70***
Numerocity	29.73***	9.14**	109.91***
Time	31.28***	102.82***	22.21***
	Non-Confidence Measure		
Fullness	21.12***	3.73	79.08***
Numerocity	48.42***	.83	191.88***
Time	15.14***	34.13***	20.34***

[a] df = 4,116

[b] df = 1,116

***p < .001

between the five equivalence interval measure means and the five non-confidence measure means is .72 for the fullness judgments, .97 for the numerocity judgments, and .92 for the time judgments. The coefficients are all highly significant,[2] indicating a substantial overlap in the variance explained by the two measures.

The overall treatment effect for each judgment type was highly significant for both the equivalence interval measures and the non-confidence measures. More important, both of these measures demonstrated the predicted pattern of uncertainty for the fullness and numerocity judgments. The pattern of means for both judgment types, along with the highly significant quadratic component in each analysis, indicated that the subjects became less uncertain about the accuracy of their fullness and numerocity judgments as the actual value of the stimulus approached either the maximum or minimum possible value. As the actual value of the stimulus approached the point mid-way between these two extremes, the subjects became increasingly uncertain about the accuracy of their judgments. This pattern is somewhat clearer for the non-confidence measure than for the equivalence interval measure, since the analyses using the latter also revealed significant linear components. These linear trends appear to be due to the asymmetry of the effect. The subjects' degree of uncertainty about the accuracy of their fullness and numerocity judgments did not decrease as much when the actual stimulus value approached the maximum possible value as when it approached the minimum possible value.

Unlike the uncertainty measures for the fullness and numerocity judgments, the uncertainty measures for the time judgments did not follow the predicted pattern. Rather, the means from both the equivalence interval measure and the non-confidence measure for the time judgments are best described by a linear trend: Subjects became more and more uncertain about the accuracy of their time estimates as the length of the stimulus increased. The quadratic components for both analyses did reach significance, but these seem to reflect primarily a ceiling effect at the higher stimulus levels. Up to a point, as the length of the stimulus increased so too did the subjects' uncertainty. Beyond this point, however, further increases in the length of the stimulus did not lead to increased uncertainty.

Judgment Accuracy

It is of further interest to examine the accuracy of the subjects' best guesses about the exact value of each stimulus. This can be done by computing the absolute difference between each subject's guess and the actual stimulus value, resulting in a judgment error score. The mean error score for each judgment is presented in Table 2. Separate repeated measures analyses of variance were computed for each of the three judgment types. The F-ratios from these analyses also are presented in Table 2.

Insert Table 2 about here

The overall treatment effect for each judgment type was highly significant. The large quadratic components of both the fullness and numerocity analyses suggest that the subjects became much more accurate in making these judgments when the actual value of the stimulus approached either the maximum or minimum possible value. As the actual value of the stimulus approached the point mid-way between these two extremes the subjects' judgments tended to become less accurate. This pattern is somewhat stronger for the numerocity judgments than for the fullness judgments, as evidenced by the significant linear trend in the latter. When the actual stimulus value approached the maximum possible value for the fullness judgment the subjects' accuracy did

Table 2

Mean Judgment Error Scores and F-Ratios from the Analyses of Variance

Judgment Type	Stimulus Level[a]					F-Ratios		
	1	2	3	4	5	Overall[b]	Linear[c]	Quadratic[c]
Fullness	1.90	5.65	9.00	9.83	5.38	25.59***	31.51***	64.79***
Numerocity	1.82	9.90	9.40	8.45	2.32	26.43***	.03	99.78***
Time	2.03	4.87	6.37	7.91	9.20	24.85***	92.22***	3.91*

Note: Larger mean values indicate greater error in judgment.

[a]The stimulus levels for the fullness judgment were 6, 24, 42, 60, and 78 marbles, respectively. The stimulus levels for the numerocity judgments were 8, 29, 50, 71, and 92 red disks, respectively. The stimulus levels for the time judgments were 6, 18, 30, 42, and 54 seconds, respectively.

[b]df = 4,326

[c]df = 1,236

*p < .05

***p < .001

not improve quite as much as when the actual stimulus value approached the minimum possible value. It is this asymmetry which apparently led to the significant linear trend. Finally, while both the linear and quadratic components of the time analysis reached significance, the linear component was clearly much stronger. In general, the subjects' time estimates became less accurate as the length of the stimulus increased.

Equivalence Interval Effectiveness

It is possible to determine how often those subjects who constructed an equivalence interval around each judgment actually enclosed the correct stimulus value, and whether this varied according to stimulus level. This can be done by assigning the subjects a 0 each time their equivalence interval enclosed the correct value, and a 1 each time it did not. The mean effectiveness score for each judgment is reported in Table 3. Separate repeated measures analyses of variance were computed for each judgment type. The F-ratios from these analyses also are presented in Table 3.

Insert Table 3 about here

As can be seen, the subjects generally constructed intervals that were too narrow. Averaging over all fifteen judgments, they failed to enclose the correct value nearly 42% of the time. More importantly, this failure to enclose the correct stimulus value varied systematically across stimulus levels. The significant quadratic component of both the fullness and numerocity analyses indicates that for these two judgment types the subjects were more likely to enclose the correct stimulus value when the actual stimulus value approached either the maximum or minimum possible value. As the actual value of the stimulus approached the point mid-way between these two extremes,

Table 3

Mean Effectiveness Score F-Ratios from the Analyses of Variance for the Equivalence Interval Measure

Judgment Type	Stimulus Level[a]					F-Ratios		
	1	2	3	4	5	Overall[b]	Linear[c]	Quadratic[c]
Fullness	.10	.33	.60	.60	.40	6.67***	11.50***	14.04***
Numerocity	.07	.63	.50	.53	.13	11.72***	.02	40.05***
Time	.27	.53	.50	.50	.57	2.28	5.21*	1.56

Note: Smaller mean values indicate greater effectiveness.

[a]The stimulus levels for the fullness judgment were 6, 24, 42, 60, and 78 marbles, respectively. The stimulus levels for the numerocity judgment were 8, 29, 50, 71, and 92 red disks, respectively. The stimulus levels for the time judgments were 6, 18, 30, 42, and 54 seconds, respectively.

[b]df = 4,116

[c]df = 1,116

$^{*}p < .05$

$^{***}p < .001$

the subjects became increasingly less likely to enclose the correct value. Again, since the linear component of the fullness analysis also reached significance, this pattern is not quite as strong as it is for the numerocity judgments. When the actual stimulus value approached the maximum possible value for the fullness judgment the subjects' ability to enclose the correct value did not improve as much as when the actual stimulus value approached the minimum possible value. Finally, the subjects' ability to enclose the correct stimulus value for the time judgments showed a slight, though significant, tendency to decrease as the length of the stimulus increased. As the stimuli became longer the subjects were less likely to enclose the correct value within the equivalence interval.

Discussion

As expected, the width of the subjects' equivalence intervals were highly correlated with their reported confidence in the accuracy of each judgment. As the subjects' confidence in the correctness of their answers decreased, the range of answers that they thought might reasonably be correct increased. It thus seems justifiable to conclude that the equivalence interval is indeed an alternative measure of uncertainty.

The findings from the present study also provide support for the hypothesis that individuals will in general be more uncertain about their judgments the further the actual stimulus value is from a clearly defined standard of comparison, or anchor point. The results based on the fullness and numerocity judgments are consistent with this prediction. As the actual stimulus value approached either the maximum or minimum possible values the subjects became less and less uncertain about the accuracy of their judgments. The results based on the time judgments, on the other hand, follow a different pattern. The subjects did seem to become less uncertain about the accuracy of their time judgments as the actual length of the time interval approached the minimum possible value. However, their uncertainty apparently did not decrease when the length of the time interval approached the maximum possible value. Rather, it tended to remain at a relatively high level.

The overall pattern of results, while not completely as predicted, can nevertheless be explained in terms of the original uncertainty hypothesis if it is assumed that some of the established maxima and minima did not provide very clear standards of comparison. For example, in retrospect it seems quite unlikely that the arbitrary 60 second maximum placed on the time intervals provided the subjects with a very good standard of comparison. Even though they were given two practice trials to help establish the 60 second interval as an anchor point at the upper end of the scale, the subjects were probably so unfamiliar with the exact duration of various time intervals in everyday life that this procedure had relatively little impact. Therefore, while 0 seconds did provide a clear anchor point for making comparisons, 60 seconds did not. If this is the case then it is not unreasonable for the subjects to be just as uncertain about the accuracy of their time judgments at the very high end of the continuum as they were at intermediate levels: In neither case did they have a satisfactory standard of comparison for making their judgments.

The subjects' uncertainty about the accuracy of their judgments at the various stimulus levels closely paralleled their actual degree of accuracy at those levels. When they were more uncertain about the accuracy of their judgments, those judgments were in fact more inaccurate. Interestingly, however, those who were given the opportunity to construct equivalence intervals around their best guesses were not able to effectively compensate for their inaccuracy in terms of being able to enclose the correct value within the interval. Even though their equivalence intervals increased in size when their best guesses were most likely to be wrong, they still failed to enclose the correct value as much as 60% of the time in some cases. These findings suggest that the subjects' level of uncertainty does not perfectly map onto their objective probability of being correct. In general, the subjects' equivalence intervals seem to indicate that they are more confident in the accuracy of their judgments than they really should be. This seems to be particularly true at the mid-points of the judgment scale, when a standard of comparison is not readily available. Similar results have been obtained by Lichtenstein, Fischhoff and Philips (1977).

These findings raise an important question. In a sense, the equivalence interval is the phenomenological counterpart of the statistical concept of a confidence interval. Yet it says little about the phenomenological level of confidence at which the subjects are operating. Is it the 95% level? The 60% level? The 40% level? Or does the level of confidence vary across individuals and situations? In order to fully understand the relationship between uncertainty and decision making behavior this question needs to be answered.

Overall, the equivalence interval technique seems to be a useful way to measure uncertainty and appears to have several advantages over other possible measures. First, it is potentially applicable in a wide variety of situations. Although the present study was concerned only with uncertainty about judgments of physical and temporal characteristics, the equivalence interval technique should work equally well for any quantitative dimension, such as uncertainty about production costs, net earnings, and industry volatility. Second, it should be possible to make specific predictions about behavior by observing whether a critical stimulus value lies inside or outside the equivalence interval. Production foremen, for example, should be much more likely to work toward a 15% production increase if this value lies within what they perceive as a reasonable range of possibilities. Finally, the equivalence interval technique provides a vehicle that can be used to further explore both the nature of uncertainty and its impact on behavior. Some work is already being done, for example, to investigate how peoples' uncertainty about various elemental aspects of the decision environment contribute to their overall decision uncertainty (e.g. Johnson, Note 2). The equivalence interval technique should prove to be quite useful in this regard.

Reference Notes

1. Larson, J. R., Jr. The effects of uncertainty about control over outcomes on preference for working conditions and information search behavior. Unpublished doctoral dissertation, Department of Psychology, University of Washington, 1977.

2. Johnson, L. C. Aggregation of uncertainty about subjective judgment. Unpublished manuscript, Department of Psychology, University of Washington, 1977.

References

Beach, L. R., Beach, B. H., Carter, W. B., & Barclay, S. Five studies of subjective equivalence. Organizational Behavior and Human Performance, 1974, 12, 351-371.

Beach, L. R., & Solak, F. Subjective judgments of acceptable error. Organizational Behavior and Human Performance, 1969, 4, 242-251.

Downey, H. K., Hellriegel, D., & Slocum, J. W., Jr. Environmental uncertainty: The construct and its application. Administrative Science Quarterly, 1975, 20, 613-629.

Downey, H. K., & Slocum, H. W., Jr. Uncertainty: Measures, research, and sources of variation. Academy of Management Journal, 1975, 18, 562-578.

Ellsberg, D. Risk, ambiguity, and the savage axioms. Quarterly Journal of Economics, 1961, 75, 643-669.

Laestadius, J. E. Tolerance for errors in intuitive mean estimation. Organizational Behavior and Human Performance, 1970, 5, 121-124.

Larson, J. R., Jr., & Mitchell, T. R. Changes in behavior following changes in control over outcomes: A theory based on responses to uncertainty. JSAS Catalog of Selected Documents in Psychology, 1977, 7, 5. (Ms. No. 1411)

Lawrence, P. R., & Lorsch, J. W. Organization and environment. Boston: Harvard Business School, Division of Research, 1967.

Lichtenstein, S. C., Fischhoff, B., & Philips, L. Calibration of probabilities: The state of the art. In H. Jungermann & G. deZeeuw (Eds.), Decision making and change in human affairs. Boston: Reidel, 1977.

Raiffa, H. Risk, ambiguity and the savage axioms: Comment. Quarterly Journal of Economics, 1961, 75, 690-694.

Stevens, S. S. On the operation known as judgment. American Scientist, 1966, 54, 385-401.

Thompson, J. D. Organizations in action. New York: McGraw-Hill, 1967.

Footnotes

[1]This research was partially supported by the Office of Naval Research Contract No. N00014-76-C-0193 (Terence R. Mitchell and Lee Roy Beach, Principal Investigators). I would like to thank Ilene Gochman for her helpful comments on an earlier draft of this manuscript.

[2]With three degrees of freedom only the correlation coefficients for the numerocity and time judgments are significant at the .05 level. This test of significance is too conservative, however, since means are being correlated instead of individual scores. The means are less influenced by measurement error and are thus more stable than are individual scores. A more appropriate test might be to use 28 degrees of freedom, based on the total number of subjects contributing to each mean. Such a test suggests that all three coefficients are highly significant, $p < .001$.

Office of Naval Research
(Code 452)
800 N. Quincy St.
Arlington, VA 22217

Director
U.S. Naval Research Laboratory
Washington, DC 20390
ATTN: Technical Info. Division

Defense Documentation Center
Building 5
Cameron Station
Alexandria, VA 22314

Library, Code 2029
U.S. Naval Research Laboratory
Washington, DC 20390

Science & Technology Division
Library of Congress
Washington, DC 20540

Psychologist
ONR Branch Office
495 Summer St.
Boston, MA 02210

Psychologist
ONR Branch Office
1030 E. Green St.
Pasadena, CA 91106

Research Psychologist
ONR Branch Office
536 S. Clark St.
Chicago, IL 60605

Director
Cybernetics Technology Office
ARPA, Room 625
1400 Wilson Blvd.
Arlington, VA 22209

Dr. H. Russell Bernard
Dept. of Sociology & Anthro.
West Virginia University
Morgantown, WV 26506

Dr. Arthur Blaiwes
Naval Training Equipment Ctr.
Orlando, FL 32813

Dr. Arie Lewin
Duke University
Duke Station
Durham, NC 27706

Dr. Lyman W. Porter
University of California
Dean, Grad. School of Admin.
Irvine, CA 92650

Dr. Paul Wall
Division of Behavioral Science
Tuskegee Institute
Tuskegee, AL 36088

Navy Personnel R & D Center
Code 01
San Diego, CA 92152

Director, Engineering Psych.
Programs, Code 455
Office of Naval Research
800 North Quincy Street
Arlington, VA 22217

Lt. Col. Henry L. Taylor, USAF
OAD(E&LS) ODDR&E
Pentagon, Rm. 3D129
Washington, DC 20301

Pers. Logistics Plans, OP987P10
Office of the Chief of Naval
Operations
Department of the Navy
Washington, DC 20350

Dr. A. L. Slafkosky
Scientific Advisor
Commandant of the Marine Corps
Code RD-1
Washington, DC 20380

Office of Naval Research
International Programs
Code 102IP
800 North Quincy Street
Arlington, VA 22217

Naval Analysis Programs, 431
Office of Naval Research
800 N. Quincy Street
Arlington, VA 22217

Operations Research Prog., 434
Office of Naval Research
800 N. Quincy Street
Arlington, VA 22217

Capt. Paul Nelson
Naval Medical R&D Command
Code 44
Naval Medical Center
Bethesda, MD 20014

Director, Behavioral Sciences
Department
Naval Medical Research Inst.
Bethesda, MD 20014

Dr. George Moeller
Head, Human Factors Engin. Br.
Submarine Medical Research Lab.
Naval Submarine Base
Groton, CT 06340

Bureau of Naval Personnel
Special Asst. for Res. Liaison
PERS-OR
Washington, DC 20370

Dr. Fred Muckler
Manned Systems Design, Code 311
Navy Personnel Research and
Development Center
San Diego, CA 92152

LCDR P. M. Curran
Human Factors Engineering Br.
Crew Systems Department
Naval Air Development Ctr.
Johnsville, Warminster, PA 18974

Dr. Alfred F. Smode
Training Analysis & Eval. Group
Naval Training Equipment Ctr.
Code N-00T
Orlando, FL 32813

Dr. Gary Poock
Operations Research Department
Naval Postgraduate School
Monterey, CA 93940

Mr. J. Barber
Headquarters DA, DAPE-PBR
Washington, DC 20546

Dr. Joseph Zeidner
Dir., Org. & Sys. Res. Lab.
U.S. Army Research Institute
5001 Eisenhower Ave.
Alexandria, VA 22304

Dr. Edgar M. Johnson
Organizations & Sys. Res. Lab.
U.S. Army Research Lab.
5001 Eisenhower Ave.
Alexandria, VA 22304

Technical Director
U.S. Army Human Engineering Labs
Aberdeen Proving Ground
Aberdeen, MD 21005

U.S. Air Force Office of
Scientific Research
Life Sciences Directorate, NL
Bolling Air Force Base
Washington, DC 20332

Dr. Donald A. Topmiller
Chief, Systems Effect. Branch
Human Engineering Division
Wright Patterson AFB, OH 45433

Lt. Col. Joseph A. Birt
Human Engineering Division
Aerospace Medical Research Lab.
Wright Patterson AFB, OH 45433

Lt. Col. John Courtright
Headquarters
AMD/RDH
Brooks AFB, Texas 78235

Dr. Jessee Orlansky
Institute for Defense Analyses
400 Army-Navy Drive
Arlington, VA 22202

Journal Supp. Abstract Service
American Psychological Assoc.
1200 17th Street, NW
Washington, DC 20036

Dr. Victor Fields
Montgomery College
Dept. of Psychology
Rockville, MD 20850

Dr. Robert R. Mackie
Human Factors Research, Inc.
Santa Barbara Research Park
6780 Cortona Drive
Goleta, CA 93017

Mr. Alan J. Pesch
Eclectech Associates, Inc.
Post Office Box 179
North Stonington, CT 06359

Dr. A. I. Siegel
Applied Psychological Services
404 East Lancaster Street
Wayne, PA 19087

Dr. W. S. Vaughan
Oceanautics, Inc.
3308 Dodge Park Road
Landover, MD 20785

Director, Human Factors Wing
Defense & Divil Institute of
Environmental Medicine
P.O. Box 2000
Downsville, Toronto, Ont., CAN.

Dr. A. D. Baddeley
Director, Applied Psych. Unit
Medical Research Council
15 Chaucer Road
Cambridge, CB2 2EF ENGLAND

Prof. Dr. Carl Graf Hoyos
Institute for Psychology
Technical University
8000 Munich - Arcisstr 21
FEDERAL REPUBLIC OF GERMANY

Dr. William A. McClelland
Human Resources Research Office
300 N. Washington Street
Alexandria, VA 22314

Major David Dianich
DSMS
Building 202
Fort Belvior, VA 22060

Dr. C. Kelly
Decisions and Designs, Inc.
Suite 600
7900 Westpark Drive
McLean, VA 22101

Dr. Paul Slovic
Oregon Research Institute
Post Office Box 3196
Eugene, OR 97403

Dr. Amos Freedy
Perceptronics, Inc.
6271 Variel Avenue
Woodland Hills, CA 91364

Dr. R. A. Howard
Stanford University
Stanford, CA 94305

Dr. Ward Edwards
Director, Social Science
Research Institute
University of South. California
Los Angeles, CA 90007

Robert G. Gough, Major, USAF
Associate Professor
Department of Economics,
Geography and Management
USAF Academy, Colorado 80840

Dr. T. Owen Jacobs
P.O. Box 3122
Ft. Leavenworth, KS 66027

Dr. Delbert M. Nebeker
Department of the Navy
Navy Personnel Res. & Dev. Ctr.
San Diego, CA 92152

Professor Ken Hammond
Department of Psychology
University of Colorado
Boulder, CO 80302

Dr. Charles Gettys
Department of Psychology
University of Oklahoma
Norman, OK 73069